

threonine glycosylation sites in polymer backbone. Fluorescence intensity was quantified along each lane of the dual-probed lectin blot, and the PNA: VVA ratio was normalized to that of the KEPAPPTP (SEQ ID NO:1)  $\times 20$  mucin; data presented as the mean and SEM from three independent experiments. (f) The fold change in PNA: VVA ratio with doubling the indicated mucin backbone size from 40 to 80 tandem repeats; data presented as the mean and SEM from three independent experiments. \*  $p < 0.05$

**[0022]** FIG. 5: Tuning Mucin Glycosylation through Cytoplasmic Tail Engineering. (a) Components and features of cell-surface mucins with synthetic 21-amino-acid transmembrane anchors (TM21) and engineered cytoplasmic motifs; native CT refers to a native cytoplasmic tail adapted from Muc1. (b) Lectin blot analysis of the indicated mucin isoforms from transiently transfected HEK293T cells to detect sialylated O-glycans by periodate oxidation and Core-I structures by PNA; blots are representative of three independent experiments. (c) PNA-lectin blot analysis of the indicated mucin isoforms before and after sialidase treatment; blots are representative of three independent experiments. (d) Top: Representative MAA and PNA lectin blot analysis (from four independent experiments) of the indicated mucin isoforms immunoprecipitated from transiently transfected HEK293T cells. Bottom: Ratiometric intensity of sialic acid to Core 1 glycan signal (MAA: PNA); data presented as the mean and SEM from four independent experiments. \*  $P < 0.05$

**[0023]** FIG. 6: Western blot analysis of MCF10A cells edited with lentivirus with native repetitive (Native\_Muc1) versus codon-scrambled Muc1 cDNAs (Muc1\_42).

**[0024]** FIG. 7: Mucins with Tunable Sizes. The sequences shown in FIG. 7 are PDTRPAPGSTAPPAHGV TSA (SEQ ID NO:8). (a) Components and features of mucin constructs with GFP reporter, native Muc1 transmembrane anchor, and codon-scrambled Muc1 tandem repeats. (b) Representative immunofluorescence images of transiently transfected HEK293T cells expressing the GFP-tagged Muc1 constructs illustrated in (a) and co-stained with PNA, anti-Muc1 antibody, and Hoechst nuclear stain (scale bar 10  $\mu$ m) from three independent experiments. (c) Components and features of mucin constructs with synthetic 21-amino-acid transmembrane anchor (TM21) and codon-scrambled Muc1 repeats. (d) Predicted molecular weight for mucin polypeptide backbone illustrated in (c). (e) Representative Western blot analysis (of three independent experiments) of TM21 constructs illustrated in (c) from extracts of transiently transfected HEK293T cells and probed with PNA lectin or anti-Muc1 antibody. (f) Representative phase-contrast images of HEK293 Ts expressed indicated constructs in (c) from three independent experiments (scale bar 100  $\mu$ m).

**[0025]** FIG. 8: Western blot Image of affinity-purified recombinant secreted mucins from FreeStyle™ 293-F cell culture media probed with anti-6 $\times$ His antibody and VVA lectin

**[0026]** FIG. 9: Cell-Surface Mucin Mutants Derived from Muc1 Tandem Repeat Sequences. The sequences shown in FIG. 9 under mMUC1 mutants (21 repeats) from top down are PDTRPAPGSTAPPAHGV TSA (SEQ ID NO:8), PDTRPAPGATAPPAHGV TSA (SEQ ID NO:5) PDTRPAPGATAPPAHGV TAA (SEQ ID NO:6) and PDARPAPGATAPPAHGV TAA (SEQ ID NO:7). (a) Components and features of mucins constructed with 21 native or engineered Muc1 repeats, GFP reporter and native Muc transmembrane

anchor. (b) Tandem repeats and predicted backbone molecular weight of native Muc1 (mMuc1) or engineered variants with single, double, or triple serine/threonine to alanine substitutions (mMuc1S, mMuc1D, or mMuc1T). (c) Representative Western and lectin blot analysis of indicated constructs in (a) from extracts of transiently transfected HEK293T cells and probed with anti-GFP antibody or co-stained with PNA, VVA and s-WGA lectins from three independent experiments. (d) Components and features of mucins constructed with 21 native or engineered Muc1 repeats and a synthetic 21-amino-acid transmembrane anchor (TM21). (e) Representative immunofluorescence images of transiently transfected HEK293T cells expressing the indicated constructs in (d) and co-stained with PNA lectin and Hoechst nuclear stain from three independent experiments (scale bar 10  $\mu$ m)

**[0027]** FIG. 10: MALDI-TOF\_MS spectra of mucin-type O-glycans as reported by Cellular O-Glycome Reporter/Amplification (CORA). HEK293T cells were transiently transfected with the indicated synthetic mucin constructs or mock vehicle. Spectra were normalized to the matrix peak at  $m/z=550$ .

**[0028]** FIG. 11: Mucins Constructed with Designer Tandem Repeats. The sequences shown in FIG. 11 are DAATPAP (SEQ ID NO:2) DAATPAPP (SEQ ID NO:3) and PPASTSAPG (SEQ ID NO:4). (a) Components and features of mucin constructs with designer tandem repeats, GFP reporter and native Muc1 transmembrane anchor. (b) Representative immunofluorescence images of transiently transfected HEK293T cells expressing the indicated GFP-tagged constructs and co-stained with PNA lectin and Hoescht nuclear stain from three independent experiments (scale bar 10  $\mu$ m).

## Part II Figures

**[0029]** FIG. 12: Engineering Biopolymer-Coated Cell Lines. A transposon-based method was used to stably integrate the DNA encoding the engineered biopolymers under a doxycycline inducible promoter. A, Schematic representation of the all-in-one vector used for producing biopolymer-coated cell lines showing key elements. For incorporation into the cellular genome, the vector includes a tetracycline responsive element (tetO), a minimal CMV promoter, the Muc1 signal sequence (Muc1 N-terminus), the tandem repeats of the biopolymer (0, 21, or 42 repeats of PDTRPAPGSTAPPAHGV TSA (SEQ ID NO:8), the transmembrane domain of Muc1 (Muc1 TM), the bicistronic green fluorescent protein reporter (IRES GFP), a EF-1 $\alpha$  promoter, the reverse tetracycline transactivator (rtTA), and a second bicistronic neomycin resistance cassette (IRES NeoR). These elements were all flanked by 5' and 3' inverted terminal repeat sequences (ITRs) required for transposon-mediated incorporation into the genome. For vector replication and production in bacteria, there was also an ampicillin resistance cassette (AmpR) and an origin of replication (ori). B, Schematic representation of membrane bound biopolymers expressed by the cells and localized to the cells surface. C, Schematic of the relative size of the extracellular domain of the engineered biopolymers designated Mucin-0, Mucin-135, and Mucin-270 for their respective length in nm. The predicted molecular weight of these proteins was 42 kDa, 81 kDa, and 120 kDa, respectively.

**[0030]** FIG. 13: Validation of Biopolymer Coatings. Expression and cell-surface localization of biopolymer coat-